

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 1 年 3 月 2 2 日
Date of Application:

出 願 番 号 特 願 2 0 0 1 - 0 8 3 0 3 8
Application Number:
[ST. 10/C]: [J P 2 0 0 1 - 0 8 3 0 3 8]

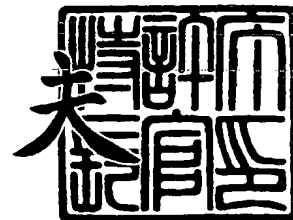
出 願 人 キヤノン株式会社
Applicant(s):

*App'n. No.: 10/653,210
Filed: September 3, 2003
Inv.: Keiichi Sakai, et al.
Title: Information Processing Apparatus And
Method, And Program*

2 0 0 3 年 9 月 1 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康



【書類名】 特許願

【整理番号】 4258010

【提出日】 平成13年 3月22日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/16

【発明の名称】 プログラム及び情報処理装置

【請求項の数】 7

【発明者】

 【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社
社内

 【氏名】 酒井 桂一

【発明者】

 【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社
社内

 【氏名】 小坂 哲夫

【特許出願人】

 【識別番号】 000001007

 【氏名又は名称】 キヤノン株式会社

【代理人】

 【識別番号】 100076428

 【弁理士】

 【氏名又は名称】 大塚 康德

 【電話番号】 03-5276-3241

【選任した代理人】

 【識別番号】 100112508

 【弁理士】

 【氏名又は名称】 高柳 司郎

 【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100115071

【弁理士】

【氏名又は名称】 大塚 康弘

【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100116894

【弁理士】

【氏名又は名称】 木村 秀二

【電話番号】 03-5276-3241

【手数料の表示】

【予納台帳番号】 003458

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0102485

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 プログラム及び情報処理装置

【特許請求の範囲】

【請求項 1】 ユーザからの音声に基づいて、ディスプレイ上に表示された複数の入力フォームのうち、ユーザがデータを入力する入力フォームを選択するために、コンピュータを、

各々の前記入力フォームの名称を記憶する記憶手段、

ユーザが発声した音声を認識する認識手段、

前記認識手段による認識の結果、ユーザが発声した音声、前記記憶手段に記憶されたいずれかの前記入力フォームの名称に該当する場合、当該名称が付された前記入力フォームをデータを入力する入力フォームとして選択する選択手段、として機能させるプログラム。

【請求項 2】 ユーザからの音声に基づいて、ディスプレイ上に表示された複数の入力フォームのうち、ユーザがデータを入力する入力フォームを選択するために、コンピュータを、

各々の前記入力フォームの前記ディスプレイ上の配置を示す配置情報を記憶する記憶手段、

ユーザが発声した音声を認識する認識手段、

前記認識手段による認識の結果、ユーザが発声した音声、前記ディスプレイ上の位置を示す場合、前記記憶手段に記憶された前記配置情報に基づいて、データを入力する前記入力フォームを選択する選択手段、として機能させるプログラム。

【請求項 3】 更に、コンピュータを、

前記選択手段により選択された前記入力フォームの表示を変更する手段、として機能させるプログラムを含むことを特徴とする請求項 1 又は 2 に記載のプログラム。

【請求項 4】 更に、コンピュータを、

前記選択手段により前記入力フォームの選択がされた場合に、音を出力する手段、として機能させるプログラムを含むことを特徴とする請求項 1 又は 2 に記載

のプログラム。

【請求項 5】 更に、コンピュータを、

前記選択手段により選択された前記入力フォームの位置を、前記ディスプレイの表示画面の中央に移動する手段、として機能させるプログラムを含むことを特徴とする請求項 1 又は 2 に記載のプログラム。

【請求項 6】 ディスプレイ上に表示された複数の入力フォームのうち、ユーザがデータを入力する入力フォームを選択する情報処理装置であって、

ユーザからの音声が入力される入力手段と、

各々の前記入力フォームの名称を記憶する記憶手段と、

ユーザが発声した音声を認識する認識手段と、

前記認識手段による認識の結果、ユーザが発声した音声、前記記憶手段に記憶されたいずれかの前記入力フォームの名称に該当する場合、当該名称が付された前記入力フォームをデータを入力する入力フォームとして選択する選択手段と

、

を備えたことを特徴とする情報処理装置。

【請求項 7】 ディスプレイ上に表示された複数の入力フォームのうち、ユーザがデータを入力する入力フォームを選択する情報処理装置であって、

ユーザからの音声が入力される入力手段と、

各々の前記入力フォームの前記ディスプレイ上の配置を示す配置情報を記憶する記憶手段と、

ユーザが発声した音声を認識する認識手段と、

前記認識手段による認識の結果、ユーザが発声した音声、前記ディスプレイ上の位置を示す場合、前記記憶手段に記憶された前記配置情報に基づいて、データを入力する前記入力フォームを選択する選択手段と、

を備えたことを特徴とする情報処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声認識技術に関し、特に、グラフィカル・ユーザインタフェース

(GUI)における音声入力に好適な音声認識技術に関する。

【0002】

【従来の技術】

GUIの入力項目(入力フォーム)へ文字等のデータを入力する場合、一般には、複数の入力フォームの中から一つを選択した上で、キーボードでデータを入力したり、マウスなどのポインティングデバイスで、提示された複数の選択候補から一つの候補を選択して入力を行なう。また、そのような入力フォームへのデータの入力に際して、音声認識技術を用いて音声入力する技術も提案されている。

【0003】

【発明が解決しようとする課題】

しかし、上記従来技術によれば、入力フォームへ音声入力をする場合に、データの输入の対象となる入力フォームをキーボードやマウスにより選択する必要がある。従って、音声入力と手操作を併用することが要求され、操作性が必ずしもよいものとはいえない。

【0004】

そこで、本発明の目的は、音声入力により入力フォームの選択を行える技術を提供することにある。

【0005】

【課題を達成するための手段】

本発明によれば、ユーザからの音声に基づいて、ディスプレイ上に表示された複数の入力フォームのうち、ユーザがデータを入力する入力フォームを選択するために、コンピュータを、

各々の前記入力フォームの名称を記憶する記憶手段、

ユーザが発声した音声を認識する認識手段、

前記認識手段による認識の結果、ユーザが発声した音声、前記記憶手段に記憶されたいずれかの前記入力フォームの名称に該当する場合、当該名称が付された前記入力フォームをデータを入力する入力フォームとして選択する選択手段、として機能させるプログラムが提供される。

【0006】

また、本発明によれば、ユーザからの音声に基づいて、ディスプレイ上に表示された複数の入力フォームのうち、ユーザがデータを入力する入力フォームを選択するために、コンピュータを、

各々の前記入力フォームの前記ディスプレイ上の配置を示す配置情報を記憶する記憶手段、

ユーザが発声した音声を認識する認識手段、

前記認識手段による認識の結果、ユーザが発声した音声、前記ディスプレイ上の位置を示す場合、前記記憶手段に記憶された前記配置情報に基づいて、データを入力する前記入力フォームを選択する選択手段、
として機能させるプログラムが提供される。

【0007】

また、本発明によれば、ディスプレイ上に表示された複数の入力フォームのうち、ユーザがデータを入力する入力フォームを選択する情報処理装置であって、

ユーザからの音声が入力される入力手段と、

各々の前記入力フォームの名称を記憶する記憶手段と、

ユーザが発声した音声を認識する認識手段と、

前記認識手段による認識の結果、ユーザが発声した音声、前記記憶手段に記憶されたいずれかの前記入力フォームの名称に該当する場合、当該名称が付された前記入力フォームをデータを入力する入力フォームとして選択する選択手段と

を備えたことを特徴とする情報処理装置が提供される。

【0008】

また、本発明によれば、ディスプレイ上に表示された複数の入力フォームのうち、ユーザがデータを入力する入力フォームを選択する情報処理装置であって、

ユーザからの音声が入力される入力手段と、

各々の前記入力フォームの前記ディスプレイ上の配置を示す配置情報を記憶する記憶手段と、

ユーザが発声した音声を認識する認識手段と、

前記認識手段による認識の結果、ユーザが発声した音声、前記ディスプレイ上の位置を示す場合、前記記憶手段に記憶された前記配置情報に基づいて、データを入力する前記入力フォームを選択する選択手段と、
を備えたことを特徴とする情報処理装置が提供される。

【0009】

【発明の実施の形態】

図1は、本発明の一実施形態に係る情報処理装置のハードウェアの構成例を示すブロック図である。情報処理装置は、G U I を表示するためのディスプレイ装置1と、数値演算・制御等の処理を行なうC P U等の中央処理装置2と、本実施形態の手順や処理に必要な一時的なデータおよびプログラム、若しくは、音声認識用の文法や音声モデルなどが格納される、ディスク装置等の外部メモリ装置若しくはR A M・R O M等の内部メモリといった記憶装置3と、ユーザが発声する音声が入力されるマイク5と、マイク5から入力された音声データをアナログ信号からデジタル信号へ変換するA/D変換器4と、を備える。これらの構成は、バスにより接続されている。

【0010】

図2は、中央処理装置2が実行する処理を示すフローチャートである。以下、本フローチャートに則して説明を行なう。このフローチャートの各ステップの操作は、プログラムとして記憶装置3に格納されており、中央処理装置2がそのプログラムを読み出して実行する。

【0011】

S1では、複数の入力フォームを含むG U I をディスプレイ装置1上に表示する。G U I の表示はH T M L のような外部データを読み込んできて表示しても、プログラムのみで表示しても構わない。

【0012】

図3は、ディスプレイ装置1上に表示される複数の入力フォームを含むG U I の表示例を示した図である。このG U I は、個人登録データの変更を想定したもので、四角形の枠が入力フォームである。例えば、入力フォーム6はID番号、入力フォーム7、9乃至13、及び、15乃至22は文字列データを入力する入力

フォームである。入力フォーム 8 及び 14 は、予め用意されたデータ（男、女、会社員等）からデータを選択するラジオボタン型の入力フォームである。ボタン 23 は、アプリケーションにデータを送信するためのボタンである。

【0013】

これらの各入力フォームの配置は、記憶装置 3 にテーブル形式で格納されている。図 4 は、各入力フォームの配置関係が記録されたテーブルの一例を示す図であり、各入力フォームに付された名称と、その配置（縦、横の位置）と、が記録されている。

【0014】

なお、本実施形態では、データの入力の対象となっている（以下、フォーカスが当たっているともいう。）入力フォームの枠線を破線で、フォーカスが当たっていない入力フォームの枠線を実線で示す。図 3 の場合、ID 番号の入力フォーム 6 の枠線が破線となっており、これにフォーカスが当たっていることを示している。

【0015】

また、図 3 に示した個人登録データの GUI は、その変更の例であり、既に変更前の登録データが存在することを想定している。そこで、図 5 に示すように、入力フォーム 6 に ID 番号を入力して、送信ボタン 23 を押下すると、図 6 に示すように、現在登録されているデータが表示されるとともに、とりあえず、自宅住所の入力フォーム 9 にフォーカスが当たる。

【0016】

次に S2 では、記憶装置 3 に格納された音声認識文法のデータを読み出す。音声認識用文法のデータには、フォーカスを当てる入力フォームを音声入力により選択するための移動認識用文法のデータと、現在フォーカスが当たっている入力フォームへのデータの音声入力を行うための項目値認識用文法のデータと、がある。移動認識用文法のデータには、「所属」や「所在地」といった入力フォームの名称を音声認識するのに必要なデータが含まれ、例えば、それらの読み方のデータが含まれる。

【0017】

なお、ここでいう文法とは、単語認識の場合、単語の表記や読みを記述した単語リスト、連続音声認識の場合 C F G（文脈自由文法）を基本としたネットワーク文法など、一般的に音声認識で用いられるものを挙げるができる。

【0018】

S 3では音声入力を開始する。ユーザーの発声した音声はマイクロフォン 5で電気信号に変換され、さらに A/D変換器 4によりデジタル信号に変換される。デジタル変換された音声信号は次の S 4の音声認識処理に用いられる。

【0019】

S 4では、入力された音声を S 2で記憶装置 3から読込んだ音声認識文法のデータにより音声認識する。この場合、移動認識用文法のデータと項目値認識用文法のデータと、のそれぞれにおいて音声認識を行う。これら 2つの文法を利用しているため、それぞれの文法から認識結果が得られるが、これを尤度など音声認識の確からしさを示す数値により比較し、より確からしさが高い結果を最終的な認識結果として選択する。

【0020】

S 5では、S 4の音声認識の結果、移動認識用文法の確からしが高かった場合に S 6に移り、項目値認識用文法の確からしが高かった場合に S 8に移る。

【0021】

S 6では、データ入力の対象となる入力フォームを選択する。認識結果として「所属」や「所在地」といった入力フォームの名称が得られた場合は、S 5からこの S 6の処理へ移行し、音声認識したその名称を、図 4に示したテーブルを参照して、該当する入力フォームを特定し、データを入力する入力フォームとして選択する。一例として、音声認識の結果が「所属」であった場合のディスプレイ装置 1の表示例を図 7に示す。

【0022】

S 7では選択確認動作を行なう。これは、選択された入力フォームをユーザに示す処理である。例えば、選択された入力フォームをフラッシュ（ある一定時間フォームの色を変える）する等して、その表示を変更したり、ビープ音などの音を鳴らしたり、或いは、選択された入力フォームが表示画面の中央に現れるよ

うに、画面をスクロールする、といった処理を行う。

【0023】

S8においては、フォーカスが当たっている入力フォームへ入力されたデータの認識結果を表示する。これは従来の技術と同様なので説明を省略する。

【0024】

このように、本実施形態では、ユーザが入力フォームの名称を発声すると、これが音声認識されて、対応する入力フォームがデータの入力対象となるので、キーボードやマウスにより入力フォームを選択する必要がない。従って、専ら音声入力だけで、GUIに対するデータ入力が可能であり、操作性が向上する。

【0025】

<他の実施形態>

上記実施形態では、入力フォームの名称がユーザにより発声された場合に、これを音声認識してデータを入力する入力フォームを選択するようにしたが、入力フォームの絶対的な位置をあらわす表現、例えば、「上から5番目」、「下から3番目」、がユーザから発声された場合に、これを音声認識してデータを入力する入力フォームを選択するようにすることもできる。

【0026】

この場合、移動認識用文法のデータには、これらの表現を音声認識するのに必要なデータが含まれ、例えば、「上から」、「下から」、「右から」、「左から」や、数字や「番目」、等を認識するデータが含まれる。

【0027】

この実施形態の場合、図2のS6におけるフォーカス移動処理においては、ユーザから発声された音声を認識し、図4に示したテーブルを参照してフォーカスを当てる入力フォームを選択する。例えば、「下から2番目」とユーザから発声された場合には、図4に示したテーブルの縦位置の最大値が11であることから、縦位置が10である電話番号の入力フォームが選択され、その位置にフォーカスを移動し、S7に移ることとなる。

【0028】

また、ユーザから、例えば、「3つ上」や「2つ下」と発声された場合には、

現在フォーカスが当たっている入力フォームからの相対的な位置の指示であると判定して入力フォームを選択するようにしてもよい。例えば、現在フォーカスが当たっている入力フォームが、「所属」の場合（図7）に、ユーザから「3つ上」と発声されると、「所属」の配置が図4のテーブルより（8、1）なので、図8に示すように、フォーカスが「FAX」（配置：（5、1））の入力フォームに当てられることとなる。

【0029】

【発明の効果】

以上説明したとおり、本発明によれば、音声入力により入力フォームの選択を行うことができる。

【図面の簡単な説明】

【図1】

本発明の一実施形態に係る情報処理装置のハードウェアの構成例を示すブロック図である。

【図2】

中央処理装置2が実行する処理を示すフローチャートである。

【図3】

ディスプレイ装置1上に表示される複数の入力フォームを含むGUIの表示例を示した図である。

【図4】

各入力フォームの配置関係が記録されたテーブルの一例を示す図である。

【図5】

ディスプレイ装置1上に表示される複数の入力フォームを含むGUIの他の表示例を示した図である。

【図6】

ディスプレイ装置1上に表示される複数の入力フォームを含むGUIの他の表示例を示した図である。

【図7】

ディスプレイ装置1上に表示される複数の入力フォームを含むGUIの他の表

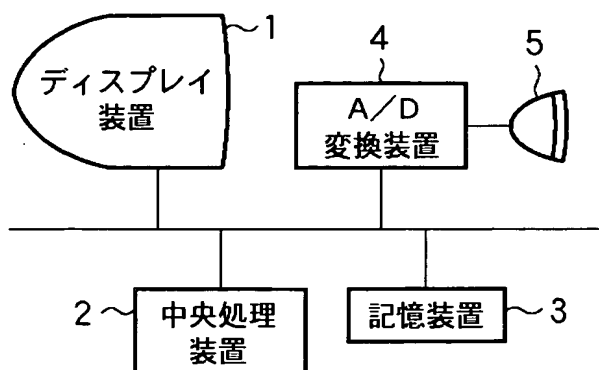
示例を示した図である。

【図 8】

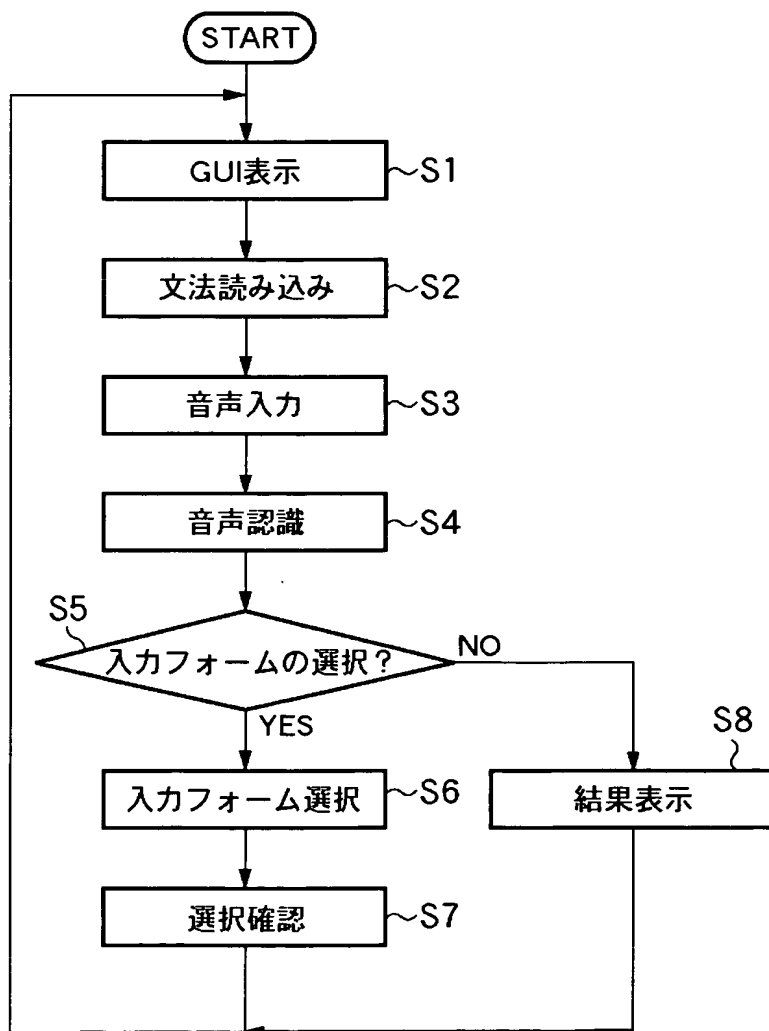
ディスプレイ装置 1 上に表示される複数の入力フォームを含む G U I の他の表示例を示した図である。

【書類名】 図面

【図 1】



【図 2】



【図 3】

個人登録データの変更

ID番号	<input type="text"/>	6	
指名	<input type="text"/>	7	性別 男○ 女○ 8
自宅住所	<input type="text"/>		9
電話番号	<input type="text"/>	10	携帯/PHS <input type="text"/> 11
FAX	<input type="text"/>	12	E-mail <input type="text"/> 13
職業	会社員○ 公務員○ 主婦○ 学生○ その他○ 14		
組織名	<input type="text"/>	15	
所属	<input type="text"/>	16	役職 <input type="text"/> 17
所在地	<input type="text"/>		18
電話番号	<input type="text"/>	19	携帯/PHS <input type="text"/> 20
FAX	<input type="text"/>	21	E-mail <input type="text"/> 22
			23 ~ <input type="button" value="送信"/>

【図 4】

入力フォームの名称	配置
ID番号	(1,1)
氏名	(2,1)
性別	(2,2)
自宅住所	(3,1)
電話番号	(4,1)
携帯/PHS	(4,2)
FAX	(5,1)
E-mail	(5,2)
職業	(6,1)
組織名	(7,1)
所属	(8,1)
役職	(8,2)
所在地	(9,1)
電話番号	(10,1)
携帯/PHS	(10,2)
FAX	(11,1)
E-mail	(11,2)

【図 5】

個人登録データの変更			
ID番号	<input type="text" value="1 2 3 4"/>	~6	
指名	<input type="text"/>	~7	性別 男○ 女○ ~8
自宅住所	<input type="text"/>		~9
電話番号	<input type="text"/>	~10	携帯/PHS <input type="text"/> ~11
FAX	<input type="text"/>	~12	E-mail <input type="text"/> ~13
職業	会社員○ 公務員○ 主婦○ 学生○ その他○ ~14		
組織名	<input type="text"/>		~15
所属	<input type="text"/>	~16	役職 <input type="text"/> ~17
所在地	<input type="text"/>		~18
電話番号	<input type="text"/>	~19	携帯/PHS <input type="text"/> ~20
FAX	<input type="text"/>	~21	E-mail <input type="text"/> ~22
			23... <input type="button" value="送信"/>

【図 6】

個人登録データの変更			
ID番号	<input type="text" value="1 2 3 4"/>	～6	
指名	<input type="text" value="鈴木一郎"/>	～7	性別 男 <input checked="" type="radio"/> 女 <input type="radio"/> ～8
自宅住所	<input type="text" value="横浜市××区○○町1-2-3"/>		～9
電話番号	<input type="text" value="045-678-9012"/>	～10	携帯/PHS <input type="text"/> ～11
FAX	<input type="text"/>	～12	E-mail <input type="text"/> ～13
職業	会社員 <input type="radio"/> 公務員 <input type="radio"/> 主婦 <input type="radio"/> 学生 <input checked="" type="radio"/> その他 <input type="radio"/> ～14		
組織名	<input type="text" value="凸凹大学"/>	～15	
所属	<input type="text" value="☆☆学部★★学科"/>	～16	役職 <input type="text"/> ～17
所在地	<input type="text" value="東京都△△区◆◆町4-5"/>		～18
電話番号	<input type="text" value="030-4567-8901"/>	～19	携帯/PHS <input type="text"/> ～20
FAX	<input type="text" value="030-4567-8902"/>	～21	E-mail <input type="text" value="suzuki@dekoboko.ac.jp"/> ～22
			23～ <input type="button" value="送信"/>

【図7】

個人登録データの変更			
ID番号	<input type="text" value="1 2 3 4"/>	~6	
指名	<input type="text" value="鈴木一郎"/>	~7	性別 男 <input checked="" type="radio"/> 女 <input type="radio"/> ~8
自宅住所	<input type="text" value="横浜市××区○○町1-2-3"/>		~9
電話番号	<input type="text" value="045-678-9012"/>	~10	携帯/PHS <input type="text"/> ~11
FAX	<input type="text"/>	~12	E-mail <input type="text"/> ~13
職業	会社員 <input type="radio"/> 公務員 <input type="radio"/> 主婦 <input type="radio"/> 学生 <input checked="" type="radio"/> その他 <input type="radio"/> ~14		
組織名	<input type="text" value="凸凹大学"/>		~15
所属	<input type="text" value="☆☆学部★★学科"/>	~16	役職 <input type="text"/> ~17
所在地	<input type="text" value="東京都△△区◆◆町4-5"/>		~18
電話番号	<input type="text" value="030-4567-8901"/>	~19	携帯/PHS <input type="text"/> ~20
FAX	<input type="text" value="030-4567-8902"/>	~21	E-mail <input type="text" value="suzuki@dekoboko.ac.jp"/> ~22
			23 ~ <input type="button" value="送信"/>

【図8】

個人登録データの変更			
ID番号	<input type="text" value="1 2 3 4"/>	～6	
指名	<input type="text" value="鈴木一郎"/>	～7	性別 男 <input checked="" type="radio"/> 女 <input type="radio"/> ～8
自宅住所	<input type="text" value="横浜市××区○○町1-2-3"/>		～9
電話番号	<input type="text" value="045-678-9012"/>	～10	携帯/PHS <input type="text"/> ～11
FAX	<input type="text"/>	～12	E-mail <input type="text"/> ～13
職業	会社員 <input type="radio"/> 公務員 <input type="radio"/> 主婦 <input type="radio"/> 学生 <input checked="" type="radio"/> その他 <input type="radio"/> ～14		
組織名	<input type="text" value="凸凹大学"/>	～15	
所属	<input type="text" value="☆☆学部★★学科"/>	～16	役職 <input type="text"/> ～17
所在地	<input type="text" value="東京都△△区◆◆町4-5"/>		～18
電話番号	<input type="text" value="030-4567-8901"/>	～19	携帯/PHS <input type="text"/> ～20
FAX	<input type="text" value="030-4567-8902"/>	～21	E-mail <input type="text" value="suzuki@dekoboko.ac.jp"/> ～22
			23～ <input type="button" value="送信"/>

【書類名】 要約書

【要約】

【課題】 音声入力により入力フォームの選択を実現すること。

【解決手段】 ユーザからの音声に基づいて、ディスプレイ上に表示された複数の入力フォームのうち、ユーザがデータを入力する入力フォームを選択するために、各々の入力フォームの名称を記憶しておき、ユーザが発声した音声を認識し（S 4）、認識の結果、ユーザが発声した音声で、記憶しておいたいずれかの入力フォームの名称に該当する場合、当該名称が付された入力フォームをデータを入力する入力フォームとして選択する（S 6）。

【選択図】 図 2

特願 2 0 0 1 - 0 8 3 0 3 8

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 1 0 0 7]

1. 変更年月日

1 9 9 0 年 8 月 3 0 日

[変更理由]

新規登録

住 所

東京都大田区下丸子 3 丁目 3 0 番 2 号

氏 名

キヤノン株式会社